Name $\qquad$ Period $\qquad$ Date $\qquad$

## AP Statistics - Unit 2 Review

1) What are the 3 principles of experimental design?

## Control, Randomization, Replication

2) What does it mean when two variables are confounded?

We cannot separate their effects on the response variable. We cannot determine what effect each variable has on the response.
3) What is the difference between subjects and individuals (or experimental units)?

Subject $=$ people $\quad$ experimental units $=$ anything else
4) What is the placebo effect?

When an experimental unit reacts to a placebo. Can be + or - reaction. The experimental unit shows a response to the placebo.
5) What is the difference between a placebo and a control?

Placebo is a dummy treatment that looks/tastes/smells/etc. the same as the real treatment.
Control is a group that we can compare our treatment too. Can be no treatment or an old treatment.
6) What is a lurking variable?

A variable that is not taken into account in our experiment, but that has an effect on the response variable.
7) We have 21 people that we need to assign to 3 different treatments (trt1, trt2, trt3). Use the section of the table of random digits below and assign the 21 people to the 3 treatment groups. List the numbers that are selected under each of the 3 headings below. Clearly explain your procedure!!

## Procedure:

- Read across the TRD below in sets of 2 digits
- Assign each person a number \#01-21.
- Ignore repeats and $\# 00,22-99$
- First 7 numbers that come up in our range are assigned to the Trt. 1, next 7 numbers that come up are assigned to Trt. 2, and whomever is left is assigned to Trt. 3.

TABLE OF RANDOM DIGITS:
$08|76| 8 \quad 1|12| 02|\quad 34| 85|9 \quad 0| 92|17| \quad 18|19| 4 \quad 4|56| 21\left|\begin{array}{llllll} & 05|07| 8 & 6|68| 13 \mid & 65|46| 1 & 5|04| 16 \mid & 99|74| 2\end{array} 0\right| 86|57|$

Sample:
Trt 1: 08, 12, 02, 17, 18, 19, 21

Trt. 2: 05, 07, 13, 15, 04, 16, 20

Trt. 3: 01, 03, 06, 09, 10, 11, 14
8) We want to test the effectiveness of a new cream designed to help healing of cuts and scrapes, against the current cream on the market (Neosporin). We do not feel the need for a placebo cream. We also want to test a new pill that is on the market that claims to speed in healing. There are 80 patients available for the experiment, of which 35 have cuts, and 45 have scrapes/abrasions.
a. What is are the factors? What are the levels of the factors?

| Factors: | Cream and Pill |
| :--- | :--- |
| Levels: | Cream: New cream and Current Cream |
|  | Pill: New Pill and Placebo |

b. What are the treatments?

New Cream and Pill
New Cream and Placebo pill
Current Cream and Pill
Current Cream and Placebo pill
c. What is the response variable?

Time it takes the wound to heal
d. What are the individuals/subjects?

Subjects: 35 patients with cuts and 45 patients with scrapes/abrasions
e. Design a completely randomized experiment, with blocking included.

9) A researcher wants to see if more expensive mattresses really give a better night's sleep than the discount mattress brand. So they recruit 110 adult volunteers to participate in their study. They will have the adults sleep on the mattresses for 10 nights and then rank their overall quality of sleep (due to the mattress only).
a. What is the explanatory variable?
mattresses
b. What are the treatments?

Expensive mattress \&discount mattress
c. What is the response variable?

Quality of sleep (ranking)
d. What are the individuals/subjects? Subjects $=\mathbf{1 1 0}$ adult volunteers
e. Design a matched pairs experiment:


* = Compare quality of sleep

10) An investigator wants to study the effects of two different fertilizers on plant growth (call them A and B). There are 20 plots available to test the fertilizers on. The investigator will measure the amount of growth by the plants after 3 months.
a. What is the explanatory variable?

Fertilizer
b. What are the treatments?

Fertilizer A and B
c. What is the response variable?

Amount of growth after 3 months
d. What are the individuals/subjects?

Individualslexperimental units $=\mathbf{2 0}$ plots of plants
e. Do you think you should use a placebo group? How about a control group?

No need for placebo since they are plants.
Yes, control group. We want to see how the plants grow naturally, and have something to compare to.
f. What are some lurking variables when it comes to plant growth?

Sunlight, water, etc.
g. It is known that the plots get different amounts of sunlight because of where they are located on the field. Some have high sun exposure, others have medium, and some have very low sun exposure. Using this information, design a block design experiment:

11) An investigator wants to study the effectiveness of two surgical procedures to correct nearsightedness. Procedure $A$ uses cuts from a scalpel and procedure B uses a laser. The data to be collected are the degrees of improvement in vision after the procedure is performed. There are 80 nearsighted people available for the experiment.
a. What is the explanatory variable? The procedure used
b. What are the treatments?

Procedure A and Procedure B
c. What is the response variable?

Degree of improvement in the vision
d. What are the individuals/subjects?

## Subjects $=80$ nearsighted people

e. Do you think you should use a placebo group? How about a control group?

No placebo or control group necessary. Neither would work. A placebo/control is not possible- the patients would know that they were getting a dummy treatment or no treatment.
f. Design a randomized comparative experiment:

g. There are two treatments. Why is it NOT ok to do a matched pairs design for this experiment?

- If you were to do a matched pairs as " $1^{\text {st }} / 2^{\text {nd" }}$ there would be no way to "reset the patients to their original vision after the first treatment so that you can then test the next procedure.
- If you were to do the matched pairs as "left/right" you would run the risk of one procedure not working well on the patient while the other one does, and then the patient could suffer because they had one eye that was good and the other is still nearsighted.

12) A study is being done to see if magnets can help relieve back pain. Participants will sleep on top of a pad that has magnets sewn into it. There are 200 people with chronic back pain that are available for the experiment.
a. We want to use a placebo in this experiment.

- How can we do this? Describe the placebo.

Instead of magnets, we would use metal pieces that are the same weight, shape, etc. as the magnets, but have no magnetic properties. We would sew them into the pads in the same way as the magnets are

- Why is a placebo necessary in this experiment?

To see if just the act of giving them a treatment (even a useless one) will produce a response (placebo effect). We need a control group- something to compare our treatment to.
b. What is the factor? What are the levels of the factor?

The magnets: Pad with magnets and Pad without magnets
C. What are the treatments?

Pad with magnets and Pad without magnets
d. What is the response variable?

Relief in back pain
e. What are the individuals/subjects?

200 people with chronic back pain
f. Design a completely randomized experiment.

g. Design a matched pairs experiment instead.
$*=$ Compare relief
from back pain

h. Which design is the better design (CRD or matched pairs)? Justify.

Matched pairs is the better design. It eliminates the lurking variables between the subjects.
13) It is known that in a specific city the chance that a person has a red hair is only 1 in 7 . A researcher wants to conduct a study to see on average how many people in Philadelphia have red hair in a sample of 5 people. Write instructions for a simulation and conduct 10 trials. Clearly label each trial and state your conclusion for the average number of people with red hair in Philadelphia.

| 5\200 | 6 |  | -505 | 55 |  |  | -3350 | $05 \$ 66$ | 55102 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 550\|0| | -61 5 | 14 | -551 | 5 | 3\|1 | $\square 430$ | 44251 | 30 | -21【3 | $3 \longdiv { 5 0 }$ |
| 94385 | 55160 | 50680 | 68443 | 95437 | 74302 | 06204 | 71004 | 76768 | 16066 | 94109 |
| 90685 | 92058 | 81744 | 99133 | 36354 | 34292 | 90092 | 21703 | 64616 | 03431 | 47610 |
| 31968 | 61593 | 36259 | 70600 | 53491 | 95542 | 78269 | 12087 | 32204 | 81177 | 30333 |


| $\#$ red <br> heads | frequency |
| :---: | :---: |
| 0 | 11111 |
| 1 | 111 |
| 2 | 111 |
| 3 | 1 |
| 4 |  |
| 5 |  |

Instructions:

- Outcomes are red hair (1/7) and not (6/7)
- Use TRD, let red hair = 0 and not red hair =1-6, Ignore 7-9.
- One trial = 5 people
- Perform 10 trials
- Response variable = \# red heads out of 5 people

Average \# red heads = 1.1 red heads per 5 people

The World Series ends when a team wins 4 games. Suppose the Phillies are in the World Series and that sports analysts consider the Phillies to have a $65 \%$ chance of winning any individual game over their opponent. We want to estimate the likelihood of the underdog (not the Phillies) winning the World Series. We also want to see how many games are played on average. Write instructions for a simulation and conduct 10 trials. Clearly label each trial and state your conclusion.

| 31968 | 61593 | $3625 \mid 9$ | 70600 | $5349 \mid 1$ | 95542 | $7826 \mid 9$ | 12087 | 32204 | $8 \mid 1177$ | 30333 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| $8 \mid 3630$ | $0602 \mid 6$ | 89308 | 94179 | $5 \mid 4907$ | $5828 \mid 0$ | 17867 | 07990 | $850 \mid 55$ | 55279 | $833 \mid 90$ |
| 37598 | 93350 | 05666 | 55402 | 87042 | 55080 | 76185 | 19947 | 79551 | 77594 | 87381 |
| 94109 | 90685 | 92058 | 81744 | 99133 | 36354 | 34292 | 90092 | 21703 | 64616 | 03431 |
| 47610 | 99430 | 44251 | 30896 | 72183 | 39856 | 94385 | 55160 | 50680 | 68443 | 95437 |


| \# games played | frequency | winner of series |
| :---: | :---: | :--- |
| 4 | II | Phils, phils, |
| 5 | IIII | Phils, phils, phils, <br> phils |
| 6 | II | Phils, Other |
| 7 | II | Phils, phils |

Average games $=\quad 54 / 10=5.4$ games on average

Instructions:

- Outcomes are Phils win (65\%) and loss (35\%)
- Use TRD, let win =00-64 and loss =65-99
- One trial = until one team (phils or opponent) gets 4 wins
- Complete 10 trials
- Response variables are \# of games and winner of series

Underdog wins $=\quad 1 / 10=10 \%$
14) Complete the following book problems:
p. 288-291 \#7, 9, 11, 13, 26, 35, 36
(for \#35 \& 36, list possible sources of bias with the sampling method listed)
\#7, 9, 11, 13: check answers in back of book
26) (a) bias = "do you think," "required"
(b) bias = "given humanity's great tradition of exploration", "do you favor", "funding"
35) (a) bias = undercoverage: not all doctors are listed in phone book

Non response: people might not answer, or hang up
(b) bias = cluster sample, leads to undercoverage of those pages that were not selected

Undercoverage: not all businesses in phone book
Non response: might not answer phone
36) (a) bias = nonresponse: people might not be home, or might not answer, or might refuse to answer Response bias: asking people to sign petition could lead to influencing the responses
(b) bias = response bias: you are interviewing them at the cafeteria. Might influence responses

Undercoverage: only one cafeteria chosen
Response bias: face to face interview
Non response

## MUTIPLE CHOICE ANSWERS:

1- D
11-B
18- D
21-B
97-E
2- B
3- C
24- D
17-C
19- B
22-D
52- C

